

///// recenzní studie / review article //////////////////////////////////////

**PRAVIDLÁ: SPOLOČNOSŤ,
JAZYK, RACIONALITA****Rules: Society, Language,
Rationality**

Abstrakt: Táto kritická štúdia sa zameriava na problematiku spolupráce a pravidiel, diskutovanú Jaroslavom Peregrinom v knihe Človek a pravidla (Praha: Dokořán 2011). Okrem extenzívneho exkurzu do osnovy Peregrinových hlavných myšlienok a argumentov budú systematicky analyzované niektoré kľúčové otázky patriace do tejto interdisciplinárnej teoretickej oblasti, pričom sa kriticky posúdia niektoré prvky Peregrinovho špecifického prístupu. Konkrétne, budem argumentovať, že Peregrinom preferovaná koncepcia pravidiel – ako propozičného poznania, že niečo by malo (nemalo) nejako byť – nemôže dobre slúžiť jeho zámerom, ak chce udržať jeho evolučný prístup k pravidlám.

Abstract: This critical review focuses on the topic of cooperation and rules discussed in the book by Jaroslav Peregrin: Človek a pravidla [Man and Rules] (Praha: Dokořán 2011). Except of an extensive excursion into the structure of Peregrin's main ideas and arguments, it aims to provide a systematic analysis of some central issues in this interdisciplinary theoretical field, while discussing some elements of Peregrin's specific approach. In particular, it will be argued that Peregrin's preferred conception of rule – as propositional knowledge to the effect that something ought to be (or ought not to be) a certain way – cannot serve well his purposes, if Peregrin is to keep his evolutionary perspective on rules.

Kľúčové slova: pravidlá; racionalita; spoločnosť; teória hier

Keywords: rules; rationality; society; game theory

LADISLAV KOREŇ

Katedra filozofie a spoločenských vied
Filosofická fakulta UHK
Rokitanského 62, 500 03 Hradec Králové

I

Od Platóna a Aristotela je v našej myšlienkovej tradícii zakorenené chápanie človeka ako rozumného tvora organizovaného vo väčších či menších spoločenstvách na báze komplexnej koordinácie činností a spolupráce, ktorej najtypickejšími formami sú jazyková komunikácia a špecializovaná deľba práce s distribúciou jej výnosov. Aristoteles v tejto súvislosti formuluje ešte jednu myšlienku hodnú pozornosti. Čo nás odlišuje od iných spoločenských tvorov, ako je napr. sociálny hmyz, je ten fakt, že ľudské skupiny sú tmelené vzájomným povedomím ich členov o tom, čo je dobré resp. zlé, či správne resp. nesprávne, ktoré je umožnené spoločne zdieľaným jazykom.¹ I keď túto ideu príliš neobjasňuje, Aristoteles mal dobrý postreh, ak to chápeme tak, že ľudské spoločenstvá sú tmelené konvenciami a normami udržiavanými a vynucovanými pomocou rôznych praktík a inštitúcií. V tomto potom zohráva zásadnú úlohu dispozícia subtilnej komunikácie informácií na jazykovom základe, pretože posúva špecificky ľudskú koexistenciu na úroveň zložitosti, ktorá nemá vo svete prírody porovnateľnú obdobu.

V knihe *Človek a pravidla* chce Jaroslav Peregrin opodstatniť v zásade veľmi podobnú myšlienku. Človek je podľa neho rozumný a ultraspoločenský tvor vládnucci jazykom, pričom kľúč k pochopeniu špecificky ľudskej spoločenskosti, racionality i jazyka – to, čo nás odlišuje od iných tvorov, ktoré majú tiež určité spoločenské, inteligenčné či komunikatívne dispozície – je v našej unikátnej schopnosti etablovať a riadiť sa regulatívnymi konvenciami a normami – *pravidlami*.² Peregrinova pozoruhodná analýza otvára čitateľovi skutočne interdisciplinárnu oblasť teoretizovania o pôvode a povahe spolupráce a pravidiel, ktorá zahŕňa príspevky z evolučnej biológie a psychológie, ekonómie, teórie hier, filozofie, a iných disciplín. Ide o čitateľsky extrémne vďačný, pritom filozoficky fundovaný úvod do problematiky, v ktorom sa kriticky reflektujú zásadné myšlienky a trendy, a nájdeme v ňom i smelý pokus o syntézu problematiky s jednotiacim motívom *pravidiel*. To je, myslím, viac než dobrý dôvod venovať Peregrinovým úvahám o pravidlách zvýšenú pozornosť a kriticky sa vysloviť k ich osnove a nosnosti.

Vznik tejto štúdie bol podporený grantom „Člověk jako normativní tvor“, GAČR P401/10/0146 2010–2012.

¹ Pozri ARISTOTELES, *Politika*. Preložil A. Kříž. Praha: Rezek 1998.

² Jaroslav PEREGRIN, *Člověk a pravidla*. Praha: Dokořán 2011.

II

Peregrin rozoberá problém pravidiel postupne z perspektívy evolúcie, mysli, jazyka a nakoniec spoločnosti a kultúry. Odrazovým mostíkom k týmto kľúčovým témam je mu však príkladný výklad problému *riadenia sa pravidlom*, ktorý svojho času opanoval analytickú filozofiu jazyka v diskusii nad kontroverznou prácou Saula Kripkeho, *Wittgenstein on Rules and Private Language*.³

Podľa Kripkeho je bežnou a intuitívnou predstavou, že mieniť resp. chápať niečo výrazom *V* spočíva v tom, že si nejakým spôsobom osvojíme pravidlo *Pr* jeho aplikácie, ktoré nás špecifickým spôsobom „vedie“ v ďalších prípadoch. To znamená, že *Pr* determinuje – avšak normatívnym, nie kauzálnym spôsobom – ako by sme *mali* resp. *nemali* použiť *V*, čím odlišuje správne od nesprávnych aplikácií výrazu *V*. Mieniť/chať niečo výrazom *V* je však viac než len vykazovať vzorec chovania v zhode s *Pr* (t.z. určitú pravidelnosť), ale zahrňuje to vždy určitý typ intencionality. Ide o to, byť si v nejakom zmysle vedomý, že *Pr* je pravidlo, ktorým sa riadime (alebo: mať zámer riadiť sa pravidlom *Pr*), a že je možné v tom uspieť i neuspieť.

Problém s touto bežnou predstavou je podľa Kripkeho v tom, že zatiaľ čo pravidlo *Pr* (napr. pravidlo *sčítania*, v prípade že $V = „+“$) typicky pokrýva neobmedzený počet prípadov, jedinec *X*, ktorý si má *Pr* osvojiť, a následne sa ním riadiť, môže byť konfrontovaný vždy len s konečným počtom prípadov, ktoré sú v zhode s *Pr*, pričom jeho kognitívne možnosti sú tiež limitované, takže *Pr* nemôže byť „uložené“ v *X*-ovej mysli ako nejaká nekončiacia tabuľka, ktorá páruje jeho vstupy s výstupmi. Keď si ďalej uvedomíme, že ľubovoľný konečný počet prípadov, ktoré sú v zhode s *Pr*, bude rovnako v zhode i s nejakým alternatívnym pravidlom *#Pr* (napríklad *pasčítania*),⁴ vyvstáva pred nami dramatická otázka: aké fakty o *X*-ovi vlastne určujú, že sa v prípade výrazu *V* skutočne riadi práve pravidlom *Pr* a nie alternatívnym pravidlom *#Pr*. Takže pri ďalšej príležitosti jeho použitia by mal *X* postupovať podľa *Pr* (viz. sčítať čísla *x* a *y*, s ktorými ešte nikdy nekalkuloval), a nie podľa *#Pr* (viz. pasčítať *x* a *y*)?

³ Pozri Saul KRIPKE, *Wittgenstein on Rules and Private Language*. Cambridge, MA: Harvard University Press 1984.

⁴ Deviantný pojem *pasčítania* môže byť definovaný napríklad takto: *pasúčet čísel x a y je rovný ich súčtu, ak x a y sú čísla, s ktorými X už kalkuloval; inak je pasúčet x a y rovný 5*. Kripkeho výklad sa opiera o aritmetické príklady, nie je však na ne obmedzený, ako dosvedčuje jeho diskusia Goodmanovských deviantných pojmov typu *modréné* (*grue*): *x je modréné = df* (i) *x je pozorované až do 1. 1. 2015, a súčasne x je zelené*, alebo (ii) *x je pozorované po 1. 1. 2015, a súčasne x je modré*.

Radikálna skeptická odpoveď je, že žiadny objektívny základ onoho druhu neexistuje. To znie na prvé počutie neuveriteľne, pretože to implikuje, že žiadne fakty o mentálnej alebo behaviorálnej histórii bežných užívateľov jazyka, ako som napríklad ja alebo vy, neurčujú, že mienime povedzme výrazom „+“ normálny pojem sčítania, a nie deviantný pojem pasčítania. Alebo, že mienime výrazom „zelený“ normálny pojem zelenosti, a nie deviantný pojem modrlenosti, a podobne. Podľa Kripkeho musí presvedčivá filozofická odpoveď na skeptikovu výzvu nielen identifikovať druh faktov o *X*-ovi, pomocou ktorých *vysvetlí*, prečo si *X* spája s *V* určitý pojem *P* a nie nejaký alternatívny pojem *#P*, ale musí ísť o taký druh faktov, ktoré *odôvodnia*, že použitia *V* na *P*-prípady sú niečím, čo *X* *mal* urobiť (pretože tak by to bolo správne), a nie sú teda len ľubovoľnými „krokmi do neznáma“, rovnako dobrými ako alternatívne použitia *V* na *#P*-prípady. Inak kolabuje intuitívna koncepcia, podľa ktorej vysvetlenie faktu, že výraz *V* znamená pre *X*-a to, čo pre neho znamená, spočíva v tom, že *X* sa drží sa určitého pravidla, ktoré si internalizoval, ktoré určuje správne resp. nesprávne aplikácie.

Kripke potom argumentuje, že splniť obidve vyššie zmienené požiadavky súčasne je neprekonateľný problém, a že takéto fakty nemožno nájsť dokonca ani vtedy, keď odhliadneme od našich kognitívnych limitácií.⁵ Ako Peregrin zmiňuje, očividne to nemôžu byť fakty o *X*-ových aplikáciách výrazu *V*. Tie sú totiž v svojej konečnosti rovnako zlučiteľné s tým, že *X* aplikoval nejaký deviantný pojem *#P*. Čo vyslovene nespomína, a čo stojí za pripomenutie, je to, že Kripke má ambíciu diskvalifikovať i iné tradičné teórie významu, či už sa odvolávajú na mentálne inštrukcie, epizódy, stavy alebo dispozície. Základná námietka proti odpovediam, ktoré sa odvolávajú na fakty o mentálnych epizódach či stavoch, je nasledujúca. Buď postulujú ďalej neredukovateľné mentálne epizódy/stavy *sui generis* a rezignujú tak na akékoľvek informatívne vysvetlenie. Alebo predpokladajú mentálne inštrukcie (formula či pravidlo *pred okom mysle*) a problém alternatívnych interpretácií sa tak odsunul len o úroveň vyššie. To však provokuje nekonečný regres *pravidiel pre správnu interpretáciu pravidiel*. Argumenty proti dispozičným teóriám zas bazirujú na tom, že aktuálne dispozície subjektu

⁵ Tým blokujeme možnú *prima facie* odpoveď, že také fakty síce sú, ibaže nie v našom epistemickom dosahu. Tým sa celý problém posúva z epistemologickej na konstitutívnu – alebo, ak chcete, metafyzickú – úroveň. Kripkeho skeptik nespochybňuje primárne, že také fakty môžeme poznať, ale že vôbec existujú. V stávke je primárne ontologický skepticizmus (nihilizmus), ktorého triviálnym dôsledkom by samozrejme bol epistemický skepticizmus. Keby totiž neboli žiadne fakty o význame požadovaného typu, potom by sme ich nemohli poznať.

sú stále len konečné, a teda zlučiteľné s deviantnými pojmami. Každopádne, z aktuálnych dispozícií možno nanajvýš vyčítať (a predikovať) ako *X* bude používať *V*, ale nie to, ako by *X* mal používať *V*.

Určitá koncepcia riadenia sa pravidlom, ako bola vykreslená vyššie, má zdá sa vážny problém. Kripke sám má tendenciu pripisovať svojmu Wittgensteinovi tzv. *skeptické riešenie*, podľa ktorého je riadenie sa pravidlom sociálna – a nie individuálna – záležitosť. Toto riešenie ďalej podporuje návrhom koncipovať sémantiku výpovedí typu *X mieni/chápe niečo výrazom V*, ako založenú na podmienkach ich sociálnej tvrditeľnosti, ktoré reflektujú *X*-ovu konformitu v užívaní výrazu *V* s praktikami jazykovej komunity. Každopádne, ak súhlasíme, že význam má normatívnu dimenziu správnosti/nesprávnosti typu riadenia sa pravidlom, potom je podľa Kripkeho i jazyk nevyhnutne sociálnou záležitosťou. Wittgensteinov slávny argument proti súkromnému jazyku sa ukazuje byť dôsledkom všeobecnejších úvah o riadení sa pravidlom.

Ďalšie peripetie tejto diskusie môžeme odložiť stranou, pretože nie sú relevantné pre posúdenie Peregrinových zámerov.⁶ Jeho hlavným dialektickým zámerom je poukázať na to, že ak jazykovú kompetenciu vidíme ako svojho druhu riadenie sa pravidlami, potom musíme vziať do úvahy nielen sociálny, ale i implicitný charakter tohto fenoménu. V opačnom prípade stroskotáme na Scylle *regularizmu* (prípád dispozičných teórií), čo je anti-intelektualistická koncepcia, ktorá nás nikdy nedostane od toho, *čo je*, k tomu, *čo by malo byť*, alebo na Charybde *regulizmu* (prípád Kantovej koncepcie *konania podľa predstavy pravidiel*), čo je naopak až príliš intelektualistická koncepcia *pravidiel pred okom mysle*, ktorá zas provokuje nekonečný regres pravidiel pre interpretáciu pravidiel.⁷

Aby sme sa dopracovali k adekvátnejšej koncepcii pravidiel, odporúča Peregrin skombinovať postrehy Wittgensteina a Sellarsa. Wittgenstein upozorňuje na to, že normatívny fenomén riadenia sa pravidlami nespočíva v tom, že subjekt je v nejakom intencionálno-kognitívnom vzťahu k ta-

⁶ Je kontroverzné, či Kripkeho skeptické riešenie podáva uspokojivú odpoveď na ním formulovaný skeptický problém. Konkrétne, je otázkou, či konformita s praktikami komunity postačuje na vysvetlenie normatívneho fenoménu správnosti/nesprávnosti, ktorý podľa Kripkeho (interpretácie Wittgensteina) charakterizuje riadenie sa pravidlom (a teda i jazykový význam resp. intencionálny obsah, pokiaľ sú explikované skrze riadenie sa pravidlom).

⁷ Užitočné rozlíšenie medzi *regularizmom* a *regulizmom* pochádza od Roberta Brandoma. Pozri Robert BRANDOM, *Making it Explicit*. Cambridge, MA: Harvard University Press 1994. Ten nasleduje myšlienky Wilfrida Sellarsa. Pozri Wilfrid SELLARS, „Some Reflections on Language Games.“ *Philosophy of Science*, roč. 21, 1954, č. 3, s. 204–228.

kému či onakému objektu, ale v poslednej inštancii je to záležitosť zhody v správaní sa a jednaní, ktorá sa zračí v aktivitách a praktikách komunity. Sellarsov komplementárny postreh je, že regulizmus i regularizmus opomínajú tretiu možnosť, ktorou je tzv. *pattern-governed behaviour* – správanie či jednanie, v ktorom je pravidlo akoby *vtelené*. Toto metaforické vyjadrenie sa dá našťastie explikovať: správanie nachádzajúce sa v priestore normatívnych postojov a praktík komunity – čo sú typicky rôzne druhy podporovania resp. sankcionovania žiaduceho resp. nežiaduceho správania sa.

Hoci už máme určitú predbežnú predstavu o *normatívnej sile* pravidiel i o tom, v čom riadenie sa nimi nespočíva, nemáme dostatočne konkrétne a pozitívne pochopenie toho, čo pravidlá sú. Pojem pravidla sa samozrejme používa rôznymi spôsobmi. Peregrin naznačuje ním preferované použitie, keď hovorí, že pravidlá bude primárne chápať *kantovsky*, ako určité obmedzenia, ktoré pre nás vytvárajú virtuálne svety slobody v kauzálnom svete prírody, podliehajúcim prírodným zákonitostiam.⁸ Prírodné zákony sa síce tiež niekedy nazývajú „pravidlami prírody“, od tých prvých sa však líšia tým, že sa nimi nemožno neriadiť, ani ich porušiť.⁹ Takto vymedzené pravidlá Peregrin ďalej odlišuje od pravidiel podriadených prírodným zákonitostiam (napr. pravidlá stavby mostu tak, aby sa nezrútil, sú podriadené zákonu gravitácie). Peregrin nasleduje Sellarsa v tom, že pravidlo je formou súčasného povedomia/poznania, že *niečo by malo nejako byť* – formou zdieľaného propozičného poznania o určitom *deontickom obmedzení*. Odložím pre túto chvíľu stranou pár kritických výhrad, ktoré mám k tomuto poslednému vymedzeniu, pretože tie nadobudnú zreteľnejšie obrysy až keď vyložím Peregrinove úvahy o úlohe pravidiel v konštitúcii ľudskej racionality, jazyka a spoločnosti. Na tomto mieste by som zmienil menšiu pripomienku. Keďže Peregrin akcentuje rolu pravidiel *qua* obmedzení, ktoré ani tak neprikazujú (ako povely), ako zakazujú, zdá sa, že relevantné propozičné poznanie by malo byť primárne typu: *že by niečo nejako nemalo byť*. Pravidlá stanovujú medze, ktoré by sa nemali prekročiť, čím súčasne vymedzujú priestory voľnosti – prípustných jednaní – v rámci týchto medzí (viz. Peregrinove virtuálne svety slobody).¹⁰

To, že pravidlom riadené aktivity sú pre náš druh charakteristické, je nekontroverzné tvrdenie. Je ťažké si predstaviť, že by rozmanité typy viac

⁸ Pozri PEREGRIN, *Člověk a pravidla*, s. 36–37.

⁹ Skôr než o pravidlách by sme tu teda mohli hovoriť o striktných pravidelnostiach, z ktorých niet výnimiek.

¹⁰ Pozri PEREGRIN, *Člověk a pravidla*, s. 34–38.

či menej kooperatívnych aktivít, ktoré nachádzame v ľudských skupinách – rodinou počnúc a štátom končiac – mohli existovať bez obmedzení typu pravidiel. Ako a prečo sa však takéto aktivity etablovali? Evolučná teória, ktorá je v dnešnej dobe paradigmou vysvetlení typu, *ako a prečo sa taký-a-taký aspekt postupne presadil vo svete prírody*, zdôrazňuje dva základné mechanizmy: *náhodné mutácie vo vlastnostiach*, ako materiál pre *prírodný výber*, v ktorom prežívajú a reprodukujú sa vo väčšej miere jedinci s takými mutáciami, ktoré im prinášajú komparatívnu výhodu v súperení o obmedzené zdroje (čokoľvek, čo vitálne podporuje ich prežitie v prostredí a replikovanie sa, ako potrava, sexuálni partneri, atď.).

Spolupráca podľa Peregrina úzko súvisí s pravidlami; ak však chceme pravidla chápať spolu s ním v kantovskom zmysle a súčasne hovoriť vážne, a nielen viac-menej metaforicky, o spolupráci napríklad v súvislosti so sociálnym hmyzom alebo šimpanzími tlupami, potom pravidlami riadené aktivity nemožno priamo stotožniť s kooperatívnymi aktivitami. Hoci sa niekedy hovorí o „pravidlách“ (spolupráce) i v súvislosti s inými sociálnymi druhmi, sotva sa tu hovorí o pravidlách v zmysle *deontických obmedzení*. Pravidlá ako deontické obmedzenia, však môžu byť práve tým špecifickým faktorom, ktorý dodal špecificky ľudskej spolupráci novú dynamiku, a tá z nej vyformovala niečo, čo nemá v prírode obdobu.

Každopádne, už samotný fenomén elementárnej spolupráce v prírode – bez nadstavby umožnenej až pravidlami – je pre evolučnú paradigmu veľkou výzvou.¹¹ Ak si vypomôžeme Peregrinovou terminológiou, problém môžeme naservírovať takto.¹² Prírodný výber medzi tzv. *replikátormi* (akékoľvek organizmy, ktoré chápeme ako jednotky prírodného výberu, ktorý je o tom, aby po sebe zanechali do ďalších generácií čo najviac svojich kópií) poháňa mechanizmus získavania výhod v prístupe k *obmedzeným zdrojom*. Avšak spolupráca si z času na čas vyžaduje, aby jedinec nezískaval/nevyužil výhodu na úkor druhých, vzniká otázka, ako sa mohli organizmy s kooperatívnymi tendenciami presadiť v prírodnom výbere. Skutočne, ako by kooperatívna dispozícia (napr. deliť sa o potravu) získaná náhodnou mutáciou alebo sériou takých mutácií, mohla replikátorom prinášať komparatívnu výhodu v prístupe ku zdrojom?

Na prvý pohľad sa môže zdať, že riešenie problému je jednoduché: spolupráca s druhými sa vypláca z dlhodobého hľadiska viac než momen-

¹¹ Nedávno bol dokonca vyhlásený časopisom *Science* za jeden z 25 najvzrušujúcejších aktuálnych vedeckých problémov. Pozri *Science*, roč. 309, 2005, č. 57312005.

¹² Pozri *ibid*, s. 46.

tálne zisky na ich úkor. Na tom nepochybne niečo je, problém je však podľa Peregrina v tom, že pokiaľ chceme vysvetliť spoluprácu v samotných jej ko-reňoch, potom nemožno predpokladať, že by replikátory mohli dohliadnuť do budúcnosti a niečo také si spočítať. Keďže evolúciu v tejto fáze poháňa logika získavania a využívania bezprostredných výhod, replikátory, ktoré by sa uspokojili s menším podielom na pomyselnom obmedzenom koláči zdrojov, by tak robili na úkor vlastnej evolučnej zdatnosti (*fitness*), a boli by vytlačené replikátormi, ktoré by túto momentálnu výhodu uvoľnených zdrojov neváhali „bez škrupúl“ využiť.

Podobne ako ďalší bádatelia v tejto oblasti, Peregrin využíva poznatky modernej teórie hier, aby problém ilustroval a analyzoval. Teória hier je často definovaná ako odnož teórie racionálnej voľby, v ktorej sa modeluje racionálne optimálne rozhodovanie v rôznych typoch interaktívnych situácií. Optimálnosť voľby aktéra – jeho najlepšia dostupná stratégia jednania z hľadiska jeho usporiadaných preferencií medzi možnými výsledkami – pritom čiastočne závisí na tom, pre akú stratégiu sa rozhodnú druhí aktéri.¹³ V súčasnosti sa však teória hier, a špeciálne *evolučná teória hier*, stále viac používa i na modelovanie situácií, ktorých aktérmi nie sú rozumné tvory, ktoré by si zámerne vyberali medzi rôznymi dostupnými stratégiami s ohľadom na efektívne uspokojovanie svojich usporiadaných preferencií. Evolučná teória hier modeluje *veľkú hru života* tým spôsobom, že uvažuje o stratégiách či vzorcoch chovania (napr. kooperatívnych a nekooperatívnych: *deliť sa/nedeliť sa o potravu; útočiť/neútočiť*, atď.) *akoby* medzi sebou súperili v prírodnom výbere, prostredníctvom hostiteľov, ktoré ich majú pevne zakódované (typicky: geneticky).¹⁴ Ak je premisou tradičnej *racionálnej* teórie hier, že racionálny aktér pri rozhodovaní maximalizuje svoje úžitky, evolučná teória hier stavia na tom, že aktérmi sú stratégie (vzorce chovania), ktoré *akoby* maximalizujú svoju evolučnú zdatnosť ako odpoveď na konkurenčné stratégie v ich prostredí. Evolučná zdatnosť stratégie je potom tým väčšia, čím väčšie je v ďalších generáciách pomerné zastúpenie jej nositeľov v populácii.

¹³ *Teória rozhodovania* je druhým odvetvím teórie racionálnej voľby, ktoré modeluje racionálne optimálnu voľbu aktéra v situáciách, keď jediným relevantným protihráčom je, tak povediac, *príroda* samotná, a stratégie iných aktérov nie sú relevantné.

¹⁴ Presnejšie povedané, *replikátor* a) vyrába svoje kópie prostredníctvom hostiteľov, a b) determinuje určitú stratégiu (vzorec chovania). Toto vymedzenie spĺňajú napríklad gény, ktoré boli do úlohy replikátorov pasované evolučnými biológmi zhruba od 60-tych rokov minulého storočia. Za popularizáciu tejto myšlienky je zodpovedný predovšetkým Richard DAWKINS, *Sobecký gen*. Praha: Mladá fronta 1998 [1976].

Peregrinov rozbor problému spolupráce a pravidiel predpokladá evolučnú teóriu hier, pretože pri výklade etablovania elementárnej spolupráce nemožno predpokladať racionálnych aktérov tradičnej teórie hier. Pre lepšie pochopenie súvislostí i odlišností medzi týmito dvomi prístupmi najskôr sústredím pozornosť na tradičnú teóriu hier, pretože v tejto súvislosti sa často tvrdilo (a stále tvrdí), že jadro problému spolupráce je obzvlášť dobre viditeľné, keď ho modelujeme v termínoch slávnej *Väzňovej dilemy* (ďalej: VD).

Táto hra je reprezentovaná strategickou tabuľkou nasledujúceho typu, v ktorej čísla od 1 do 4 vyjadrujú úžitky (nazývané *payoffs*) aktérov A a B pre 4 možné výsledné stavy, dané kombináciami dvoch dostupných stratégií *spolupracovať* a *nespolupracovať*, medzi ktorými si aktéri môžu vyberať. Prvé číslo vyjadruje úžitky aktéra A, zatiaľ čo druhé vyjadruje úžitky aktéra B, a ordinálna škála 1–4 nám hovorí len toľko, že číslo 4 charakterizuje stav vecí *x*, ktorý aktér preferuje pred stavom vecí *y*, ktorý má priradené číslo 3, a ten aktér preferuje pred stavom vecí *z*, ktorý má priradené číslo 2, ktorý zas aktér preferuje pred stavom vecí *w*, ktorý má priradené číslo 1.¹⁵

¹⁵ Abstraktné úžitky možno v rôznych kontextoch aplikácie tohto modelu interpretovať napr. monetárne, ako subjektívne blaho aktéra, či reprodukčný úspech. To ale neznamená, že úžitky možno priamočiaro identifikovať s týmito špecifickými kategóriami. Z formálneho hľadiska racionálny aktér maximalizuje *úžitkovú funkciu* (*utility function*). To znamená len toľko, že optimálna voľba jednaní (stratégie) efektívne uspokojuje aktérovo konzistentne usporiadané preferencie medzi rôznymi výslednými stavmi vecí (*outcomes*). *Konzistentnosť preferencií* má dve zásadné vlastnosti: *tranzitívnosť*: ak (*X* preferuje *x* pred *y*), a súčasne (*X* preferuje *y* pred *z*), potom (*X* preferuje *x* pred *z*); a *úplnosť*: pre každý pár výsledných stavov *x* a *y* platí, že (*X* preferuje *x* pred *y*), alebo (*X* preferuje *y* pred *x*), alebo (*X* je indiferentný medzi *x* a *y*). Funkcia užitočnosti môže byť *ordinálna* alebo *kardinálna*. Prvá zoradí výsledné stavy od najpreferovanejšieho až po najmenej preferovaný, ale nehovorí vôbec nič o tom, ako veľmi ich aktér chce, zatiaľ čo druhá dokáže reprezentovať ich relatívnu váhu z pohľadu aktéra. Ak predpokladáme ordinálnu užitočnosť, je pre aktéra optimálne vybrať si jednanie, ktoré povedie k jeho najviac preferovanému výslednému stavu. Ak predpokladáme kardinálnu užitočnosť, je pre aktéra optimálne maximalizovať tzv. *priemerný očakávaný úžitok*. Základná myšlienka je jednoduchá. Uvažujme len 3 výsledné stavy s_1, s_2, s_3 , a určité kardinálne hodnoty $U(s_1) = 10, U(s_2) = 6, a U(s_3) = 1$, ktoré reprezentujú ich relatívnu váhu z pohľadu aktéra *X*. Majme nasledujúce pravdepodobnosti, že jednanie j_i – ktoré si *X* môže vybrať s pomedzi dostupných jednaní j_1, \dots, j_n – povedie k stavom $s_1, s_2, a s_3$: $Pr(s_1/j_1) = 0.2; Pr(s_2/j_1) = 0.3; Pr(s_3/j_1) = 0.5$. Priemerný očakávaný úžitok jednaní j_i vypočítame nasledovne: $U(j_1) = (U(s_1) \times Pr(s_1/j_1)) + (U(s_2) \times Pr(s_2/j_1)) + (U(s_3) \times Pr(s_3/j_1)) = (10 \times 0.2) + (6 \times 0.3) + (1 \times 0.5)$. Keď sa analogicky spočítajú očakávané úžitky ostatných jednaní z množiny $\{j_1, \dots, j_n\}$, maximalizácia priemerného očakávaného úžitku znamená toľko, že *X* si vyberie to jednanie z danej množiny, ktoré má najvyšší priemerný očakávaný úžitok. Dobrý, filozoficky orientovaný úvod do problematiky podáva napr. Gerald GAUS, *On Philosophy, Politics, and Economics*. Belmont: Wadsworth Publishing 2007; alebo Joseph HEATH, *Following the Rules: Practical Reasoning and Deontic Constraint*. Oxford: Oxford University Press 2008.

| | B <i>Spolupracovať</i> | B <i>Nespolupracovať</i> |
|-----------------------------|---------------------------|-----------------------------|
| A <i>Spolupracovať</i> | 3; 3 | 1; 4 |
| A <i>Nespolupracovať</i> | 4; 1 | 2; 2 |

Figura 1: Vázňova dilema

Táto strategická tabuľka zachycuje fakt, že kooperatívna stratégia síce prináša aktérovi potenciálne nezanedbateľný úžitok, avšak len v prípade, keď i druhá strana opláca rovnakou mincou. Inak spolupracujúci aktér poskytuje výhodu väčšieho podielu na pomyselnom koláči zdrojov nespôlupracujúcej strane. Tzv. *striktne dominantnou stratégiou* v interaktívnych situáciách s touto štruktúrou je *nespolupracovať*. Čokoľvek totiž urobí druhý aktér, prvý aktér je na tom vždy lepšie, keď s ním odmietne spolupracovať (to isté platí, symetricky, z pohľadu druhého aktéra). Jediné stabilné riešenie tejto hry – tzv. *Nashovo ekvilibrium* (rovnováha) – je potom kombinácia stratégií (*nespolupracovať-nespolupracovať*).¹⁶ Optimálne rozhodnutie z individuálneho hľadiska je teda nespôlupracovať. Toto uvažovanie však aktérov dovedie do výsledného stavu, ktorý je pre obidvoch horší než keby spolupracovali (podelili by sa o zdroje, neútočili by na seba, a podobne). Je pritom dôležité si všimnúť, že VD nie je *hrou s nulovým súčtom* (všeobecnejšie: *hrou s konštantným súčtom*), v ktorej by sa hralo, obrazne povedané, o podiel na fixnom koláči, kde jeden aktér môže získať väčší podiel vždy len na úkor druhého (čo prvý získa, o to druhý príde). Zatiaľ čo hra s nulovým (konštantným) súčtom je o čistom konflikte, vo VD je okrem motívu kon-

¹⁶ Nashovo ekvilibrium pre hru je ľubovoľná kombinácia stratégií jej hráčov, pre ktorú platí: z hľadiska každého hráča je jeho stratégia najlepšou možnou odpoveďou na stratégie ostatných hráčov. Inak povedané: *žiadny hráč si nemôže polepšiť, ak sa jednostranne odchýli od svojej stratégie, v prípade, že sa ostatní hráči držia svojich stratégií*. Ekonóm a matematik John Nash na začiatku 50. rokov 20. st. dokázal, že každá hra s konečným počtom ťahov má aspoň jedno Nashovo ekvilibrium, ktoré však môže existovať i v tzv. *miešaných stratégiách*, ktoré predstavujú určité *randomizácie* voľby aktéra, podľa pravdepodobností, s ktorými vyberá tú alebo onú čistú stratégiu. Napríklad v elementárnej hre typu *kameň-papier-nožnice* je takouto randomizovanou stratégiou, ktoré je Nashovým ekvilibrium: hrať s rovnakou pravdepodobnosťou kameň, papier i nožnice (a keďže sa hrá typicky na viac kôl, ide o stratégiu: hrať tretinu času kameň, tretinu času papier, a tretinu času nožnice).

fliktu prítomný i motív spolupráce. Zohľadňuje sa v nej fakt, že spojenie síl v spolupráci môže byť podstatne produktívnejšie než nespolupráca.

Často možno počuť názor, že VD vystihuje jadro problému spolupráce alebo dokonca morálky, ktorý je treba filozoficky riešiť tak, že sa ukáže, že je v nejakom zmysle *racionálne* spolupracovať v situáciách typu VD, hoci sa to tak nejaví z pohľadu *individuálnej racionality*, a nemožno to teda zachytiť štandardnými prostriedkami teórie racionálnej voľby. Tu by sme však mali byť opatrní. Čo presne má znamenať, že je racionálne spolupracovať v situáciách typu VD?

Ak nechceme riskovať, že sa dopustíme kategóriálnej chyby prisúdením racionality akémusi abstraktnému kolektívnemu hrdinovi (aktérovi), potom to, čo majú kritici teórie racionálnej voľby v tejto súvislosti pravdepodobne na mysli, je to, že je racionálne smerovať rozhodnutia ku globálne výhodnejšiemu kooperatívnemu výsledku, i keď tento výsledok nie je Nashovým ekvilibriom VD. Kritik tu implicitne predpokladá, že individuálna racionalita, vymedzená ako efektívne uspokojovanie usporiadaných preferencií, je *de facto* o tom, že racionálny aktér je sebec, ktorý bez škrupúl sleduje vlastné záujmy a ciele bez ohľadu na záujmy a ciele druhých. Ak je to tak, teória racionálnej voľby aplikuje jednorozmerný model racionálneho aktéra, ktorý má problém riešiť sociálne dilemy inak než nekooperatívne. Podľa kritika je však nepopierateľným faktom skúsenosti, podporeným v poslednej dobe navyše i experimentálne, že ľudia sú pripravení spolupracovať v situáciách typu VD, čo len potvrdzuje neadekvátnosť modelu individuálnej racionality. Keď však aktéra doplníme o iné, charakteristicky ľudské rozmery, ako sú altruistické a prosociálne sentimenty či internalizované normy (*férovosti*, *zlatého pravidla*, a podobne), potom uvidíme, že spolupráca je celkom dobre možná.

Na margo tejto kritiky teórie racionálnej voľby je treba povedať dve veci. Po prvé, model individuálnej racionality efektívneho uspokojovania preferencií, nemusí byť adekvátny pre všetky teoretické účely. V tomto prípade však kritik dezinterpretuje aparát racionálnej voľby. Vôbec nič nebráni tomu, aby funkcia užitočnosti formálne zachytila ľubovoľné konzistentne usporiadané preferencie aktéra, a to vrátane altruistických preferencií alebo preferencií, ktoré zohľadňujú záujmy druhých aktérov.¹⁷ Po druhé, je treba si uvedomiť, že ak aktér skutočne preferuje nespolupracovať, keď druhý aktér spolupracuje, i keď druhý nespolupracuje – a to je práve zachytené v strate-

¹⁷ K tomu pozri dobrú diskusiu v GAUS, *On Philosophy, Politics and Economics*, alebo HEATH, *Following the Rules*.

gickej tabuľke VD – potom je pre neho len a len racionálne nespolupracovať. Situácia sa samozrejme mení v prípade, že aktér preferuje spolupracovať, keď druhý spolupracuje (a nespolupracovať, keď nespolupracuje). Potom sa však aktérove usporiadané preferencie nezhodujú so strategickou tabuľkou VD, ale s inou hrou, ktorá je známa ako *hra na istotu* (*assurance game*).

| | B <i>Spolupracovať</i> | B <i>Nespolupracovať</i> |
|-----------------------------|---------------------------|-----------------------------|
| A <i>Spolupracovať</i> | 4; 4 | 1; 3 |
| A <i>Nespolupracovať</i> | 3; 1 | 2; 2 |

Figura 2: *Hra na istotu* (so zárukou)

Je dobré si hneď ujasniť, že všetko, čo bolo práve povedané, nemusí spochybňovať, že určitá *prima facie* tenzia individuálneho a spoločného záujmu je charakteristická pre morálne problémy a sociálne dilemy rôzneho druhu, ani to, že sociálne konvencie a normy typicky požadujú zohľadniť záujmy druhých (skupiny), i keby to malo byť na úkor individuálnych záujmov (prípadne, že sú prostriedkom zmierenia individuálnych záujmov s kolektívnym).¹⁸ Zdá sa ale, že regulatívne konvencie a normy môžu byť „riešením“ dilemy typu VD skôr v tom zmysle, že zmenia samotné podmienky situácie (hry), v rámci ktorých sa prevážia pomyselné misky váh v prospech globálneho či kolektívneho záujmu,¹⁹ a chránia tak aktérov od toho, aby sa príliš často stretávali v interaktívnych situáciách typu VD alebo podobných *sociálnych pascách*.²⁰ Nemusíme spochybňovať ani to, že v sebe prechová-

¹⁸ Opatrnejšiu formuláciu volím preto, že niektorí autori argumentujú, že sociálne normy a morálka nie sú v konflikte s individuálnym záujmom, ale zohľadňujú napr. dlhodobý záujem individua.

¹⁹ Pozri napr. Cristina BICCHIERI, „Social Norms.“ In: *Stanford Encyclopedia of Philosophy* [online]. Zalta, E. (ed.). Dostupné na <<http://plato.stanford.edu/entries/social-norms/>> [cit. 27. 11. 2011]. posledná verzia 2011; Kenneth BINMORE, *Natural Justice*. Oxford: Oxford University Press 2005; Bryan Skyrms, *The Stag Hunt and the Evolution of Social Contract*. Cambridge: Cambridge University Press 2004.

²⁰ Možná analógia sa často hľadá u Hobbesa. Podľa jednej možnej interpretácie, Hobbesov *prírodný stav* – v ktorom si každý jedinec môže robiť to, čo chce, podľa vlastného subjektívneho súdu, pretože neexistuje uznávaná autorita, ktorá by efektívne riešila spory a vynucovala poriadok – kolabuje *od vojny všetkých proti všetkým*, lebo dominantnou

vame i altruistické a prosociálne sentimenty, ktoré by mohli mať podobne vyvažujúci efekt. Pokiaľ sa však tieto faktory premietajú do našich preferencií v interaktívnych situáciách, potom by sa to malo premietnuť i do ich formálnej reprezentácie, ktorá je daná strategickou tabuľkou (všeobecne: funkciou užitočnosti).²¹

Každopádne, zdá sa, že v interaktívnych situáciách so štruktúrou jednoduchej VD sú vyhliadky pre spoluprácu racionálnych aktérov biedne. Normy alebo prosociálne sentimenty by síce situáciu mohli meniť v prospech spolupráce, pokiaľ ale chceme spolu s Peregrinom vysvetliť samotný vznik elementárnej spolupráce, ťažko ich môžeme predpokladať, pretože i oni volajú po evolučnom vysvetlení. Bolo už ostatne naznačené, že celý model tradičnej teórie hier je práve z tohto dôvodu nevhodný pre Peregrinov projekt: pri výklade etablovania elementárnej spolupráce nemožno predpokladať racionálnych aktérov tradičnej teórie hier, pretože a) spolupráca existuje už na úrovni organizmov, ktoré zďaleka nie sú racionálne v požadovanom zmysle,²² a b) racionalitu samotnú chce Peregrin vysvetliť skrze pravidlá, ktoré podľa neho úzko súvisia s kooperatívnymi aktivitami tým, že im typicky stanovujú, ľudovo povedané, určité mantinely.²³

stratégiou v podmienkach súperenia o obmedzené zdroje a vzájomnej nedôvery je nekooperatívna stratégia *útočiť (skôr než zaútočí druhý)*. Túto podľa Hobbesa neznesiteľnú situáciu nie je možné „vyriešiť“ len tým, že aktéri nahliadnu globálnu výhodnosť spolupráce (neútočenia). Je z nej možné sa dostať len tak, že sa zmenia jej podmienky tak, že cena za útočenie (trest) bude v transformovanej situácii tak hrozivá, že si každý dva krát rozmyslí, či bude útočiť a koristiť, či nebude dodržiavať dohody a sľuby, atď. Podľa Hobbesa nestačí len dohoda na pravidlách spoločenskej koexistencie a kooperácie (*rules of justice*), ale vyžaduje si to etablovanie – či už dohodou alebo silou – neobmedzenej autority, ktorá bude fungovať ako odstrašujúca moc, ktorá takéto pravidlá vynucuje. Pozri Thomas HOBBS, *Leviathan*. Curley, E. (ed.). Indianapolis: Hackett Publishing Company 1994. Porovnaj ale úvod E. Curleyho (*ibid.*, s. xxvi-vii), kde sa tento výklad problematizuje v tom zmysle, že i Hobbes pripúšťa, že v prirodzenom stave môže byť za určitých okolností rozumné spolupracovať, a síce v prípadoch, keď druhá strana začala spolupracou (splnila svoju časť v dohode/*covenant*). Také situácie však odpovedajú skôr hre na istotu, či *lov jeleňa* (viz. ďalšia diskusia). Pozri tiež SKYRMS, *The Stag Hunt and the Evolution of Social Contract*.

²¹ Z tohto dôvodu sú tiež diskutabilné experimenty, ktoré majú preukázať, že ľudia sú prekvapivo ochotní spolupracovať vo VD, a jej podobných situáciách (sociálnych dilemách). Dalo by sa totiž namietnuť, že aktéri sú buď iracionálni, prípadne, že takéto experimentálne situácie nemedelujú VD, pretože ich relevantné preferencie sú usporiadané spôsobom, ktorý nevyhovuje strategickej matrici VD (napr. v dôsledku toho, že aktéri už do nich vstupujú s internalizovanými sociálnymi normami podporujúcimi prosociálne, kooperatívne jednanie). Toto by však bolo na dlhšiu debatu, do ktorej sa tu nemôžem púšťať.

²² PEREGRIN, *Človek a pravidla*, s. 47.

²³ *Ibid.*, s. 42, 60.

Je však situácia podstatne odlišná, ak predpokladáme spolu s evolučnou teóriou hier, že aktérmi sú skôr samotné behaviorálne stratégie (vzorce) – skrze hostiteľské organizmy, ktoré ich majú pevne dané (*hard-wired*) – a nie racionálni aktéri? Nie, v prípade jednoduchej VD je situácia podobne pesimistická:

Potrebujeme vysvetliť, jak se mohli rozšířit „spolupracující“ tvorové, když se zdá, že museli vstupovat do střetů odpovídajících Věžňovu dilematu s tvory „nespolupracujícími“ a být jimi vytlačováni.²⁴

Skutočne, jednorazové stretnutia organizmov s pevne danou kooperačnou stratégiou (*neútočiť a podeliť sa o zdroje*) a organizmov s pevne danou nekooperačnou stratégiou (*útočiť a nepodeliť sa o zdroje*) by vychádzali fatálne v neprospech kooperačných organizmov. Pokiaľ by bola veľká hra života skutočne o tomto, prírodný výber by ju riešil v prospech nespolupracujúcich organizmov. To je však dostatočný dôvod na zamyslenie sa, pretože život v prírode je o spolupráci minimálne v rovnakej miere ako o súperení. Ak teda chceme spoluprácu vysvetliť z evolučnej perspektívy, mali by sme predpokladať, že aktéri veľkej hry života existujú v priestore interaktívnych situácií, ktorých štruktúru je treba modelovať úplne inou hrou než je jednoduchá VD.

Ponúkajú sa napríklad rôzne *koordinatívne hry*, v ktorých individuálne záujmy hráčov nie sú striktné protichodné, ale viac či menej sa prekrývajú. V poslednom čase sa napr. stále viac diskutujú hry typu *lov jeleňa* (*stag hunt*). Tá sa odvodzuje od Rousseauovej pasáže z *Rozpravy o pôvode nerovnosti medzi ľuďmi*,²⁵ kde Rousseau v súvislosti s primitívnym lovom hovorí, že i keď je úlovok veľkej koristi (jeleňa) z kolektívneho hladiska preferovanou možnosťou, vyžaduje si koordináciu aktivít, ktorá je dosť krehká, pokiaľ sú jednotliví lovci ochotní opustiť svoju funkciu v koordinovanej činnosti, a pri vzniknutej príležitosti loviť uháňať menšiu ale pravdepodobnejšie dostupnú korisť (zajaca), o ktorú sa okrem toho nemusia s nikým deliť.

²⁴ *Ibid.*, s. 48.

²⁵ Jean-Jacques ROUSSEAU, „O původu nerovnosti mezi lidmi.“ In: *Rozpravy*. Praha: Svoboda 1978, s. 75–144.

| | B (druhý lovcí) <i>loviť jeleňa</i> | B (druhý lovcí) <i>loviť zajaca</i> |
|--------------------------|--|--|
| A <i>loviť jeleňa</i> | 3; 3 | 1; 2 |
| A <i>loviť zajaca</i> | 2; 1 | 2; 2 |

Figura 3: *Lov jeleňa*

Podobne ako v prípade VD, *lov jeleňa* len exemplifikuje množinu typovo podobných interaktívnych situácií so štruktúrou danou strategickou tabuľkou. Na rozdiel od VD však v tejto hre nie je žiadna stratégia dominantná, a má hneď dve Nashove ekvilibriá v čistých stratégiách: (*loviť jeleňa*; *loviť jeleňa*) a (*loviť zajaca*; *loviť zajaca*). Takéto situácie sa na prvý pohľad javia viac naklonené spolupráci než VD, pretože kombinácia kooperatívnych stratégií (*loviť jeleňa*; *loviť jeleňa*) je Nashovým ekvilibriom hry, ku ktorému by mohla eventuálne konvergovať, a je navyše z globálneho hľadiska výhodnejšia než nekooperatívne ekvilibrium. Problém je, že kooperatívne ekvilibrium je síce dominantné z hľadiska globálnej výhodnosti, nekooperatívne ekvilibrium má ale navrch z hľadiska rizika (garantuje vyššie minimum).²⁶ Vo VD sa rieši problém, či investovať určité náklady do potenciálne obojstranne prospešnej spolupráce, pritom ale vážne riskovať, že druhý odmietne spolupracovať, a využije bez oplátky jeho investíciu vo svoj prospech? Pri *love jeleňa* stojí otázka trochu inak: investovať náklady do potenciálne výhodnej spolupráce s druhými, alebo postupovať na vlastnú päsť pri vzniknutej príležitosti, s menším potenciálnym ziskom, ale i s menším rizikom? I keby všetci lovci vzorne spolupracovali, pravdepodobnosť, že ulovia väčšiu korisť, môže byť menšia, než že jedinec svojpomocne uloví menšiu korisť. Riziko spojené so spoluprácou je každopádne vyššie, pretože jej úspech závisí na koordinovanom úsilí všetkých aktérov. Každý aktér však má stimul nespoločnosť, keď sa v jeho blízkosti objaví dostupnejšia korisť. Nekooperujúci riskuje menej: po prvé, jeho korisť je dostupnejšia; po druhé, úspech pri jej *love* nezávisí na príspevku druhých. Zdá sa teda, že všetko nasvedčuje skôr tomu, že *lov jeleňa* bude spieť skôr do nekooperatívneho ekvilibria, a nie je jasné, ako môže byť táto hra lepším modelom vzniku spolupráce, než jednoduchá VD.

²⁶ K tomu pozri SKYRMS, *The Stag Hunt and the Evolution of Social Contract*.

Základný problém kooperácie pri love jeleňa je v *dôvere*. Keďže v *love jeleňa* je výhodné spolupracovať v prípade, že druhý spolupracuje, a nespôlupracovať, v prípade, že druhý nespôlupracuje, potom, keby mohli jeden druhému dôverovať, že budú spolupracovať, konvergovali by do kooperatívneho ekvilibria. Ak si však aktér myslí, že druhý môže rovnako dobre spolupracovať ako nespôlupracovať, potom má dobrý dôvod neriskovať spoluprácu. Spoľahlivá signalizácia a komunikácia zámerov by podľa niektorých mohla nakloniť misky váh v prospech kooperatívneho ekvilibria,²⁷ čo by mohlo poukazovať na zaujímavé korelácie medzi evolúciou spolupráce a komunikácie, a na úlohu, ktorú by v evolúcii spolupráce smerom k jej komplexnejším formám mohol zohrať vývin stále subtilnejších komunikačných systémov, končiac jazykom. Iná možnosť, ktorá sa skúma, je, že interaktívne situácie nemusia byť tak náhodné, ako sa štandardne zvykli modelovať: jedinci s kooperatívnymi sklonmi majú tendenciu vyhľadávať si im podobné typy, a vyhýbať sa nespôlupracujúcim, čo by spelo k vytváraniu homogénnych skupín jedincov s podobnými (kooperatívnymi) sklonmi, čiastočne izolovanými od nespôlupracujúcich. Keď k tomu ešte pridáme dispoziciu napodobovať a učiť sa od úspešných vzorov v okolí, potom vyhliadky pre úspešné šírenie *lovcov jeleňov* môžu byť o poznanie svetlejšie.²⁸

Posledné úvahy nám naznačujú ešte inú možnosť. Situácia sa totiž dramaticky zmení, ak predpokladáme, že sa organizmy (s ich zakódovanými stratégiami) stretávajú v situáciách typu VD nie raz, ale opakovane. V tom prípade sa zdajú byť vyhliadky pre šírenie kooperácie podstatne ružovejšie, pretože nad replikátormi sa vznáša *tieň budúcnosti*: u opakovaného interagovania vzniká možnosť oplácať nespôluprácu druhej strany „trestaním“, spôsobujúcim straty, s ktorými treba rátať. Experimenty na počítačových modeloch, ktoré prebiehajú už od 80-tych rokov minulého storočia, ukazujú, že v takých podmienkach sa dokážu výrazne presadiť *milé* stratégie. Tie začínajú spoluprácu a potom kombinujú určitým spôsobom kooperatívne ťahy s trestaním.²⁹ Najznámejšia z nich – tzv. *ako ty mne, tak ja tebe* (*tit for tat*) – inštruuje nositeľov nasledovne: a) začni spoluprácou; avšak b)

²⁷ Niekedy sa však pod *lovom jeleňa* uvádza i hra, ktorá má trochu inú štruktúru: ten, kto loví zajaca, keď druhý loví jeleňa, má vyšší úžitok = 3, zatiaľ čo spolupracujúci má stále úžitok = 1. V takom prípade by ani komunikácia nemusela smerovať aktérov ku kooperatívnemu ekvilibriu, pretože pre každého z hráčov by bolo výhodné presvedčiť druhého, aby spolupracoval (lovil jeleňa), či už on sám plánuje spolupracovať alebo nie. Táto hra je však len verziou *hry na istotu*.

²⁸ Pozri SKYRMS, *The Stag Hunt and the Evolution of Social Contract*.

²⁹ Pozri Robert AXELROD, *The Evolution of Cooperation*. New York: Basic Books 1984.

v prípade, že druhý s tebou v n -tej konfrontácii nespolupracuje, oplat' mu v ďalšej rovnakou mincou; buď ale zhovievavý, a c) začni pri ďalšej príležitosti opäť kooperatívne.³⁰ Táto veľmi dobre čitateľná recipročná stratégia je jedným z Nashových ekvilibríí opakovanej VD pre dvoch hráčov. Ak si teraz na chvíľu opäť predstavíme, že aktéri môžu vyberať i striedať stratégie, a že obidvaja postupujú podľa tejto recipročnej stratégie, potom nikto z nich nemá dôvod sa od nej jednostranne odkloniť (nespolupracovať), pretože by si nepolepšil (za predpokladu, že sa hrá na neurčitý počet kôl).³¹

Treba jedným dychom dodať, že existuje mnoho iných, i podstatne menej milých stratégií, ktoré môžu byť rovnako stabilným riešením hry, alebo jej rôznych variantov, ktoré sa formulovali, aby adekvátnejšie vystihli dynamiku hry života. Toto pozorovanie je známe ako *ľudový teorém* (*folk theorem*).³² Ten hovorí, neformálne povedané, že ak je daná určitá stabilná, dostatočne vysoká pravdepodobnosť, že hra sa bude opakuvať (nech už sa pred tým opakovala n krát), potom každý výsledný stav, ktorý je pre každého z hráčov aspoň tak dobrý, ako jeho tzv. *minimaximálny bod* (v tom sa každý z hráčov snaží držať druhého na jeho minime), môže byť Nashovým ekviliбриom opakovanej hry s neurčitým počtom opakovaní. Dôvodom v podstate je, že hráčom je obojstranne známe, že druhého možno potrestať za odchýlenie sa od danej stratégie. To okrem iného predpokladá, že hráči majú starosť o svoju budúcnosť, a že sú o sebe dobre informovaní. Takto sa i stratégia naklonená spolupráci môže stať ekviliбриom opakovanej VD, a to z toho jednoduchého dôvodu, že strach z budúcich trestov preváži bezprostredné zisky plynúce z nespolupráce. Okrem *tit-for-tat*, iná kooperatívna stratégia s potenciálom ekvilibria, pomenovaná *Grim*, inštruuje aktéra, aby začal spolupracou, ale nespolupracovanie už potom stále, a bez odpúšťania,

³⁰ Táto recipročná stratégia bola evolučnými biológmi diskutovaná už od začiatku sedemdesiatych rokov dvadsiateho storočia ako základ mechanizmu tzv. recipročného altruizmu. Ten má údajne vysvetliť spoluprácu medzi nepríbuznými jedincami, ktorá je síce nepopierateľným faktom v prírode, nebolo ju však možné adekvátne vysvetliť štandardne rozoznávaným mechanizmom príbuzenského altruizmu. Pozri Robert TRIVERS, „The Evolution of Reciprocal Altruism.“ *Quarterly Review of Biology*, roč. 46, 1971, č. 1, s. 35–56.

³¹ Inak by sa totiž dalo pomocou tzv. *spätnej indukcie* (*backward induction*) dokázať, že dominantnou stratégiou, ku ktorej bude opakovaná VD s *určitým* počtom kôl konvergovať, je opäť *nespolupracovať*. Pozri napr. GAUS, *On Philosophy, Politics and Economics*, alebo sofistikovanejší výklad v Martin HOLLIS – Robert SUGDEN, „Rationality in Action.“ *Mind*, roč. 102, 1993, č. 405, s. 1–35.

³² Pretože na neho prišlo nezávisle na sebe viacero bádateľov – prvý zrejme John Nash.

oplácal len trestaním (odmietaním spolupracovať).³³ Toto pozorovanie nakoniec nie je žiadnou novinkou. Už David Hume upozornil na to, že práve takýmto mechanizmom môže vzniknúť spontánna koordinácia činností a spolupráca, a to i bez vonkajšej donucujúcej sily.³⁴

V takýchto experimentoch sa ale v prvom rade skúmalo, za akých podmienok sa môže stratégia, špeciálne kooperatívna, evolučne presadiť v prostredí, v ktorom sú zastúpené konkurenčné (nekooperatívne) stratégie, a aká stratégia môže byť tzv. *evolučne stabilnou stratégiou*.³⁵ Zhruba povedané: S je evolučne stabilnou stratégiou, ak distribúcia S v populácii P je taká, že žiadna nová *mutantná* stratégia #S sa nemôže výraznejšie presadiť v P na úkor S – akýkoľvek pomerný nárast #S v P bude časom kompenzovaný opäť v prospech S. Časom sa ukázalo, že počiatočný optimizmus s naivnou recipročnou stratégiou *tit-for-tat* nebol celkom na mieste, pretože jej evolučný úspech (schopnosť presadiť sa v populácii) závisí od toho, aké typy stratégií sa vyskytujú, v ako početnej populácii, a v akom pomernom zastúpení. Signifikantné je, že evolučne stabilnou stratégiou typicky nie je jedna stratégia, ale tzv. *zmiešaná stratégia*. V tej sa hrajú (sú zastúpené) rôzne stratégie s frekvenciou (pomerným zastúpením), ktorá(é) odpovedá ich výskytu v momente stabilizácie populácie, pričom určité nastavenie podmienok favorizuje iné miešané stratégie, v ktorých majú väčšie slovo stratégie, ktoré sú menej milé než *tit-for-tat*.³⁶

³³ Iná zaujímavá stratégia, ktorá môže tiež byť stabilným riešením, a ktorá je príznačne pomenovaná *tat-for-tit*, zas inštruuje nositeľa od začiatku nespoločnosť, a začať spolupracovať hneď potom, čo protihráč odmietne spolupracovať. K tomu pozri Matt RIDLEY, *The Origins of Virtue*. Londýn: Viking (Penguin Books) 1996, alebo Michael NOWAK, *SuperCooperators: We Need Each Other to Succeed*. Chicago: Simon and Schuster 2011, kde sa možno dozvedieť viac o najnovšom výskume v tejto oblasti.

³⁴ Pozri David HUME, *A Treatise of Human Nature*. L. A. Selby-Bigge (ed.), revised by P. H. Nidditch. Oxford: Clarendon Press 1975.

³⁵ Ide o určité precizovanie Nashovho ekvilibria pre dynamické evolučné hry; keďže už nemôžeme hovoriť o aktéroch, ktorí si môžu vyberať a striedať stratégie, hovoríme o prírodnom výbere medzi stratégiami na základe ich zdatnosti či reprodukčného úspechu: čím je stratégia úspešnejšia v súperení o zdroje v danej generácii, tým viac bude mať zástupcov v ďalšej generácii, až eventuálne bude v populácii dominovať.

³⁶ Napr. v Axelrodovom počítačovom experimente tohto druhu prežilo z počiatočných 63 stratégií, ktoré spolu po dvojiciach hrali, nakoniec len 6, pričom *tit-for-tat* mala v populácii o niečo málo viac, než šestinový podiel. Videné prizmou evolučnej teórie hier, víťazom celého „turnaja“ bola zmiešaná stratégia 6 rôznych stratégií, z ktorých každá sa hrala s určitou frekvenciou. Okrem toho, *tit-for-tat* nie je evolučne stabilná, ale len *evolučne neutrálna stratégia*: do populácie, v ktorej sú len jej zástupcovia, síce nemôžu úspešne preniknúť nekooperatívne stratégie, ale môže sa v nej dariť rovnako dobre napr. jednoduchej kooperatívnej stratégii *vždy spolupracuj*.

Je pomerne široká zhoda na tom, že vznik a šírenie kooperácie a regulatívnych noriem možno evolučne (biologicky a/alebo kultúrne) vysvetliť i za pomoci aparátu teórie hier. Toto teritórium sa však momentálne nachádza v stave tak intenzívneho bádania, že je predčasné hovoriť o niečom viac, než hypotézach vo fáze experimentálneho testovania. Niet preto divu, že ani u Peregrina nenájdeme čitateľ uspokojivé vysvetlenie evolúcie spolupráce, či pravidlami riadených kooperatívnych činností. Ostáva nezodpovedanou otázkou, ako vôbec mohli kooperatívne replikátory prekonať *počiatočnú bariéru*, a presadiť sa v populácii, v ktorej dominovali nekooperatívne stratégie, keď „prvé ťahy“ v hre života sú očividne v ich neprospech.

Peregrin nepredstiera, že disponuje definitívnymi odpoveďami, a otvorene upozorňuje na to, že kooperatívne stratégie typu *tit-for-tat* predpokladajú nie celkom triviálne kognitívne schopnosti, u ktorých nie je jasné, či ich môžeme u elementárnych replikátorov predpokladať (schopnosť evidovať s kým sa stretli, a ako sa zachoval, aby mohli patrične oplácať), rovnako ako *metastratégie*, ktoré by mohli organizmy inštruovať k tomu, aby sa vyhýbali nekooperujúcim protihráčom. Sympatické na jeho prístupe je to, že sa nebojí formulovať smelé hypotézy, ktoré konzistentne sleduje v rôznych aspektoch. Jednou takouto hypotézou je, že zárodok spolupráce a pravidlami riadených kooperatívnych aktivít je v určitých komplexných vzorcoch (stratégiách) správania. Jedinci totiž môžu mať zakódovaný nielen vzorec chovania, ale i tendenciu vnucovať druhým tento vzorec správania, či dokonca tendenciu vnucovať druhým, aby tento vzorec ďalej vnucovali druhým.

Podľa Peregrina je práve existencia takéhoto *metavzorca* charakteristická pre pravidlami riadené aktivity, ktoré majú odlišovať ľudské skupiny od elementárnejších sociálnych foriem. Toto dobre ladí s inou vecou, ktorú pred tým zmieňuje, a síce, že do dynamiky hry života môže výrazným spôsobom prehovoriť donucovací faktor *trestania* jedincov, ktorí nevykazujú kooperatívne vzorce správania. A hoci to Peregrin explicitne nezmiňuje, z jeho diskusie je zrejmé, že môžeme uvažovať i o faktore *metatrestania*, keď sa trestajú nielen nespolupracujúci, ale i jedinci, ktorí ich netrestajú (trestajú sa i jedinci, ktorí nespolupracujú na trestaní nespolupracujúcich).

Niekedy sa v tejto súvislosti hovorí o *altruistickom trestaní*, pretože tí, ktorí trestajú nespolupracujúcich, tak nezriedka robia na vlastné trovy, ktoré nemusia byť návratné.³⁷ Jednou intenzívne skúmanou hypotézou mo-

³⁷ K tomu pozri napr. Herbert GINTIS – Samuel BOWLES – Robert BOYD – Ernst FEHR, „Explaining Altruistic Behavior in Humans.“ *Evolution and Human Behavior*, roč. 24, 2003, č. 3, s. 153–172; Herbert GINTIS – Joseph HENRICH – Samuel BOWLES – Robert BOYD –

mentálne je, že tento fenomén *silnej (strong) reciprocity* nemožno redukovat' na príbuzenský, recipročný, či nepriamy (*indirect*) altruizmus (ten zdôrazňuje aspekt reputácie, ktorý vstupuje do hry, keď aktérov pri interakcii pozorujú iní potenciálni partneri v ďalších interakciách podobného typu, ktorí zaznamenávajú ich správanie). Jeho zástancovia tvrdia, že silný altruizmus vo všeobecnosti nemožno vysvetliť teóriami „osvieteného“ individuálneho záujmu (v dlhodobjšom horizonte), ale vyžaduje si jedincov, ktorí sú emocionálne patrične už uspokojení, a nachádzajú záľubenie v kooperácii a vzájomnej pomoci, a pociťujú *prosociálne emócie*, ako je súcitiť, zahanbenie alebo vina, keď druhí jedinci, alebo oni sami, nespolupracujú či ubližujú iným členom skupiny. To by potom podľa nich poukazovalo na dôležitý fakt, že špecificky ľudskú, sociálnymi normami tmelenú koexistenciu v skupinách, nemožno vysvetliť na základe rozšíreného predpokladu o prirodzenej ľudskej motivácii poháňanej individuálnym záujmom akokoľvek osvieteného časovým horizontom.

III

Nech je to už s touto prastarou otázkou ľudskej prirodzenosti a koreňov morálky akokoľvek, dostávame sa tu k zásadnému momentu. Zdá sa, že Peregrinove vysvetlenia sa v tomto bode začínajú zamotávať do nebezpečne úzkeho kruhu. Ako sme videli, jeho hlavným cieľom je ukázať, ako môžu kantovsky chápané pravidlá, ako deontické obmedzenia, zohrávať zásadnú rolu v ľudskej sociálnosti, racionalite i jazyky, a byť tým faktorom, ktorý nás odlišuje od iných druhov. Sám však upozorňuje na dve veci. Po prvé, pravidlám podliehajúcu aktivitu chápe ako formu spolupráce, ktorá požaduje na aktérovi investíciu, ktorá čosi stojí, a je pritom výhodná len za predpokladu reciprocity druhých, a je viac než pravidelnosťou či koordináciou v chovaní, pretože je založená na zdieľanom povedomí v rámci určitej skupiny, že *niečo by nejako malo (nemalo) byť*. Po druhé, i keď vždy existuje riziko *čiernych pasažierov* parazitujúcich na kooperácii a sociálnych normách, ich etablovanie a stabilitu by sme mohli ľahšie vysvetliť, ak by sme už mohli predpokladať dostatočne rozumných tvorov, ktorí by si dokázali spočítať, že spolupráca s druhými a normované aktivity sú pre nich vzájomne výhodné (v dlhodobjšom horizonte), a koordinovali by k nim ich aktivity. Lenže to podľa Peregrina práve nemôžeme predpokladať, pretože racionalita a jej

Ernst FEHR, „Strong Reciprocity and the Roots of Human Morality.“ *Social Justice Research*, roč. 21, 2008, č. 2, s. 241–253.

mentálno-kognitívne piliere predpokladajú pravidlá, respektíve predpokladajú jazyk, ktorý je konštituovaný pravidlami.

Jeden by sa však mohol diviť: na jednej strane má byť pravidlo formou zdieľaného poznania, že *niečo by nejako malo/nemalo byť*, a predpokladá teda zrejme určitú mentálno-kognitívnu kapacitu, ktorá je už pripravená pre propozičné poznanie, na druhej strane sa však hovorí, že propozičné poznanie, a iné stavy s propozičným obsahom, predpokladajú pravidlá, pretože predpokladajú jazyk, ktorý je pravidlami konštituovaný.

Aby sme si závažnosť situácie objasnili, všimnime si faktor zdieľaného (súčasného) povedomia o tom, čo by malo resp. nemalo byť. Ten vyzdvihol už Aristoteles, a v tej či onej podobe je prítomný i v moderných teóriách sociálnych noriem anticipovaných Humovým výkladom konvencií. Dôležitým elementom v týchto teóriách je,³⁸ že aktéri majú *podmienečnú preferenciu* vykazovať určitú konformitu v správaní, *pokiaľ* i dostatočné množstvo druhých vykazuje túto konformitu, a súčasne veria či očakávajú, že druhí budú vykazovať tento vzorec. Podľa Peregrina však propozičné poznanie či mentálne stavy typu presvedčení či prianí vyžadujú existenciu robustného logického priestoru, ktorý sa konštituuje pravidlami používania jazykových výrazov. Presnejšie, utvára sa až keď komunikačné praktiky dosiahnu takú úroveň

³⁸ K tomu pozri BICCHIERI, „Social Norms.“ Priekopnícke práce v tejto oblasti sú Thomas C. SCHELLING, *The Strategy of Conflict*. New York: Oxford University Press 1963; a David LEWIS, *Convention*. Cambridge, MA: Harvard University Press 1969. V rámci tejto tradície Bicchieri rozlišuje konvencie a normy, kde prvé umožňujú vybrať ekvilibrium v problémoch koordinácie typu *Na ktorej strane cesty riadiť auto?*, ktoré je výhodné pre všetkých (nikto nemá dôvod sa odchýliť, pretože je spoločným záujmom všetkých vyhnúť sa zrážkam), zatiaľ čo normy síce tiež riešia problém výberu ekvilibria, ale nemusia prinášať bezprostrednú výhodu. Často práve naopak, jedinec má individuálny stimul sa odchýliť, a získať tak výhodu čierneho pasažiera na úkor „poctivcov“. Konvencie sa potom mohli vyvinúť ako nezamýšľané riešenia určitých koordinatívnych problémov v rámci menších skupín s opakovanými a nenáhodnými interakciami členov; sociálne normy by mohli byť podobne nezamýšľanou odpoveďou skupiny na problémy opakovaných hier s miešaným motívmi, kde sa záujmy aktérov prekrývajú len čiastočne – majú koordinatívny i konfliktný aspekt – ako je opakovaná VD, hra *verejné statky* (*public goods* – verzia VD pre *n* hráčov), alebo tzv. hra *dôvery* (*trust game*). A v súlade s tým, čo som naznačoval v rozbere VD, Bicchieri špekuluje, že normy nevyberajú ekvilibriá priamo pre sociálne dilemy, ale transformujú ich na koordinatívne problémy, u ktorých potom riešia problém výberu medzi ekvilibriami, typicky, v prospech kooperatívnych a sociálne výhodných ekvilibrií. Ďalšia vec, ktorá stojí za zmienku, je, že i keď si konvencie vyžadujú masovejšiu konformitu než normy, ich udržanie sa v skupine nie je natoľko problematické, pretože sa výrazne eliminuje motivácia čierneho pasažiera. Sociálne normy sa naopak môžu v skupine udržiavať, i v prípade, že nie sú masovo dodržiavané, vďaka tomu, že sú v priestore normatívnych praktík a očakávaní: t. z. očakávaní, že druhý od nás očakáva konformitu, a prípadne si ju na nás vynuocovať.

zložitosti, že bude možné hovoriť o zmysluplných jednotkách typu viet, ktoré sú vo vzťahoch zlučiteľnosti a nezlučiteľnosti, a (materiálneho) vyplývania.

Prezradím hneď na začiatku, že Peregrinovým riešením avizovaného potenciálneho kruhu vo vysvetľovaní regulatívnych noriem je metóda tzv. *bootstrappingu*, ktorá predpokladá, ak tomu dobre rozumiem, že dva aspekty sa môžu súčasne postupne vyvíjať, s tým, že jeden sa v určitej fáze vývoja opiera o druhý, a naopak. Pravidlá sa o rozum (myslel) do určitej miery opierajú, súčasne mu (jej) však pomáhajú na svet. Aby túto myšlienku explikoval, Peregrin najskôr predstavuje model proto-mysle ako *emulátoru*. Ten v sebe akumuluje akési asociácie medzi 1) pokynmi smerom z centra organizmu do jeho periférie (motoriky), 2) senzorickými vnemami z periférie (vrátane tých, čo sú dôsledkom pokynov), a 3) stavmi „príťažlivosti“ či „odpudivosti“ (proto-emóciami), ktoré v nich vnemy prípadne vyvolávajú. Organizmus potom dokáže takéto asociácie využívať tak, že konfrontuje svoje aktuálne pokyny s podobnými pokynmi uloženými v emulátore, a pomocou porovnávania aktuálnych vnemov s uloženými vnemami (asociovanými s uloženým pokynom), koriguje svoju motoriku vzhľadom na svoje ciele (dosiahnutie potravy, vyhnutie sa prekážke, a podobne).

Organizmus takto nadobúda nielen schopnosť učiť sa so skúsenosti (štýlom *pokus-omyl*), ale i reprezentovať si situácie (zárodok presvedčení) a ciele (zárodok prání) v ich neprítomnosti (zárodok intencionality), a skúšať si „nanečisto“ určité aktivity (bezbolestný *pokus-omyl*), a na základe toho dynamicky prispôbovať svoje behaviorálne stratégie prostrediu. Tento mechanizmus sa ďalej mohol v evolúcii zdokonaľovať, v prípade ľudskej mysle však nastáva určitý kvalitatívny prielom, ktorý Peregrin vidí ako prechod od (signalizačného) *kódu zvieraťa* k propozičnému *kódu človeka*. Jeho príkladom je rozdiel medzi praktickým (*know how*) a propozičným poznaním (*know that*): prvé je určené na riešenie konkrétnych praktických problémov a možno ho neproblematicky pripisovať i iným živočíchom, zatiaľ čo druhé nie je viazané na konkrétne praktické problémy, a s jeho pripisovaním iným druhom živočíchov je to problematickejšie. Peregrin tu má zrejme na mysli, že propozičné mentálne stavy – ako sú poznatky, presvedčenia či prania – možno neproblematicky pripisovať len nám podobným tvorom, ktorí vykazujú určitý komplexný vzorec správania sa, ktorý je v samotnom základe jazykový. Aby sa z organizmu, ktorý si reprezentuje prostredie a ciele, stal rozumný aktér, ktorého motivujú presvedčenia a prania, ako *dôvody* jeho jednania, musí do hry vstúpiť kód človeka.³⁹

³⁹ Pozri PEREGRIN, *Človek a pravidla*, s. 84, 92.

Logika Peregrinovej koncepcie je nasledovná. Myslenie síce nie je internalizovanou rečou, ako si to predstavovali poniektorí behavioristi, ale materiál, vo forme propozičných obsahov (pripisovateľných „že“-vetami resp. „aby“-vetami), mu dodáva práve jazyk. Až s konštitúciou dostatočne robustného *logického priestoru výpovedí*, ktoré sú v paradigmatických vzťahoch *zlúčiteľnosti*, *nezlúčiteľnosti* a *materiálneho vyplývania*, sa konštituuujú propozičné obsahy. Tie, tak povediac, „zhmotňujú“ role výpovedí v holistickej sieti týchto vzťahov (inferenčné role).⁴⁰ Peregrin to pritom podáva tak, že z hľadiska jazyka je ústrednou jazykovou hrou *tvrdenia*, ktorá je zas vnorená v hru *na udávanie a požadovanie dôvodov*, pretože (a) kto niečo tvrdí, berie na seba určitý záväzok zdôvodniť svoje tvrdenie, v prípade, že sa spochybní, a (b) spochybniť niečo predpokladá možnosť poprieť to. Na to však už potrebujeme určité výpovedné jednotky, ktoré sú vo vzťahoch odôvodňovania (premisy a záver), a odporovania si, či nezlučiteľnosti. Inak povedané: potrebujeme už robustný logický priestor výpovedí, ktorý je podľa Peregrina vytýčený

[...] námi, užívateľi jazyka, v podobe pravidiel, jimiž se řídíme při používání slov. (Například to, že *Blíží se tygr* je možné udat jako důvod pro *Hrozí nebezpečí*, je dáno určitými pravidly, která se konstituovala v rámci procesu zavádění slov jako *tygr*, *nebezpečí* atd.). Jsou to tedy tato pravidla, která generují logické vztahy vyplývání a neslučitelnosti.⁴¹

Peregrinova inferencialistická koncepcia jazyka – zaujímavá a kontroverzná súčasne – je farbisto vyličená v kapitole „Pravidla a Jazyk“. Postup, ktorý začína jazykovou hrou tvrdenia, v rámci hry na udávanie a požadovanie dôvodov a pátra po jej predpokladoch, je podľa mňa plodný z perspektívy teórie významu, keď nás nezaujíma ani tak to, ako sa náš jazyk vyvinul do plnokrvnej podoby, ako to, čo je pre túto podobu charakteristické čo do komplexnosti jej pragmatických, syntaktických a sémantických aspektov.⁴² Z perspektívy vývoja jazyka sa mi ale zdá, že to, čo Peregrin hovorí v tejto kapitole o konštitúcii robustného logického priestoru výpovedí – i to, čo hovorí o funkcii logických výrazov ako prostriedkov explikácie pôvodne implicitných vzťahov zlučiteľnosti, nezlučiteľnosti a vyplývania medzi vý-

⁴⁰ *Ibid.*, s. 81–82.

⁴¹ *Ibid.*, s. 82–83.

⁴² Podobne ako známa Davidsonova teória významu – pozri napr. Donald DAVIDSON, „The Structure and Content of Truth.“ *Journal of Philosophy*, roč. 87, 1990, č. 6, s. 279–328 – založená na teórii radikálnej interpretácie, ktorá má, podobne ako Brandomova (BRANDOM, *Making it Explicit*), alebo Peregrinova koncepcia, výrazne pragmatický aspekt.

povedami – sa celkom dobre zlučuje s tým, že argumentačná štruktúra je super-štruktúrou, ktorá neutvára, ale naopak predpokladá určitý robustný logický priestor výpovedí, ako ho charakterizuje sám Peregrin.

Zdá sa mi napríklad pravdepodobnejšia špekulácia, že ľudské skupiny mohli najskôr rozvinúť pomerne komplexný systém komunikačných praktík, ktoré mohli utvoriť robustný logický priestor i bez toho, aby bola etablovaná normatívna jazyková hra na tvrdenie, ako ju Peregrin charakterizuje, ktorú je možné chápať napríklad tak, že vyrástla z elementárnejších foriem jazykovej komunikácie, zameranej primárne na informovanie, varovanie, a podobne (viz. napr. zaujímavý model *konšpiračného šepotu* Krebsa a Dawkinsa, o ktorom Peregrin referuje).⁴³ Jedna možnosť, ktorá sa ponúka, je, že sociálna prax tvrdenia sa etablovala ako odpoveď ľudských skupín – citlivých na dezinformácie formou nespoľahlivosti či klamaní – na potrebu „bezpečnej informačnej zmeny“ (*tvrditeľ* ako ten, kto ručí druhým za informáciu, ktorú im komunikuje).

Vráťme sa konečne k hlavnej veci, ktorá už nadobudla, ako dúfam, jasnejšie obrysy. Peregrinov prístup obsahuje sympaticky smelý pokus ukázať, že pravidlá sú tým, čo nás, ako *normatívnych tvorov*, odlišuje od iných druhov, pretože riadenie sa pravidlami je spoločným prvkom špecificky ľudskej racionality, jazyka, i spoločenskosti. Otázka znie: dokáže Peregrin skutočne prelomiť avizovaný začarovaný kruh: *pravidlá sa o myseľ opierajú, súčasne sa však rozum (myseľ) opiera o pravidlá, pretože sa opiera o jazyk, ktorý sa opiera o pravidlá?*

Spomeňme si, po prvé, že pravidlá chce chápať v užšom zmysle ako deontické obmedzenia – typu *niečo by nejako malo/nemalo by byť* – ktorých sú si členovia skupiny určitým spôsobom súčasne vedomí (viz. súčasné/zdieľané povedomie o tom, že by niečo nejako malo/nemalo byť). Dokonca i keď sa riadia pravidlom len implicitne, sú v sociálnom priestore normatívnych postojov a očakávaní zračiach sa v typicky korektívnych praktikách, ktoré sa ich týkajú. Konkrétne, u Peregrina opakovane nájdeme v rôznych metamorfózach nasledujúci motív:

[J]e to práve schopnosť takového poznávania v podobe *vědění-že*, na čem stavi pravidla.⁴⁴

⁴³ Pozri PEREGRIN, *Člověk a pravidla*, s. 106–111; Richard DAWKINS – John R. KREBS, „Animal Signals: Information or Manipulation.“ In: KREBS, J. R. – DAVIES, N. B. (eds.), *Behavioral Ecology: An Evolutionary Approach*. Oxford: Blackwell 1976, s. 380–402.

⁴⁴ PEREGRIN, *Člověk a pravidla*, s. 63.

Za druhé, *vedenie-že* a iné propozičné stavy majú byť produktom prechodu od signalizačného kódu zvierata k propozičnému kódu človeka, ten však predpokladá jazyk:

Z tohoto hľadiska není na našom jazyku tak podstatné to, jak rozsáhlým souborem slov disponujeme, jako to, že vykazuje onu logickou strukturu, která otevírá prostor pro propozice a pro *vědění-že*.⁴⁵

Za tretie, jazyk sám je konštituovaný pravidlami:

Ony logické vztahy vyplývání a neslučitelnosti, které jsou takto za konstituci propozic zodpovědné, jsou, jak podrobně uvidíme v následující kapitole, vytyčeny námi, uživateli jazyka, v podobě pravidel, jimiž se řídíme při používání slov.⁴⁶

Kruh sa nám v tomto bode nebezpečne uzatvára, a nie je na prvý pohľad úplne jasné, ako nám z neho môže pomôcť Peregrinom zmieňovaná metóda *bootstrappingu*. Ak to totiž chápeme doslovne, potom propozičné poznatky predpokladajú propozičný kód, ten však predpokladá jazyk, ktorý zas predpokladá pravidlá, ktoré sú sami propozičnými poznatkami – a tak ďalej dokola!

Kde je teda pes zakopaný? Podľa môjho názoru v tom, že Peregrin chce od jeho robustného pojmu *pravidla* niekedy až príliš veľa teoretickej práce. Nechce nám totiž povedať len to, že pravidlá normatívne výrazne utvárajú ľudský rozum, jazyk a *ultrapoločenskosť*, ale chce povedať aj niečo podstatné o tom, ako evolúcia pravidiel v rámci ľudských skupín transformovala elementárnejšie mentálno-kognitívne, komunikačné a sociálno-kooperatívne štruktúry na úroveň komplexnosti, ktorá nemá v prírode obdobu. Zdá sa mi však, že pokiaľ chce Peregrin držať tento sympaticky evolučný prístup, mal by pripustiť, že to, čo v niektorých úvahách nazýva *pravidlami*, nie sú v skutočnosti plnokrvné pravidlá (*qua* deontické obmedzenia), ale akési proto-pravidlá (podobne ako emulátor je len proto-myslou, či konšpiračný šepot je len proto-jazykom).

Summa Summarum: Peregrinova analýza fenoménu pravidiel čitateľa rozhodne odmení, a jeho úvahy o pravidlách a spolupráci sú v tom najlepšom zmysle slova „extrémne stimulujúce“, už preto, že otvárajú priestor pre filozofické debaty o problematike, ktorá je v našom prostredí zatiaľ poľom neoraným.

⁴⁵ *Ibid.*, s. 85.

⁴⁶ *Ibid.*, s. 82.